Appl. Math. Lett. Vol. 5, No. 4, pp. 29–33, 1992
Printed in Great Britain.

0893-9659/92 $5.00 + 0.00

FACTORIZATION OF POLYNOMIALS USING NILPOTENT JORDAN BLOCKS

THOMAS J. LAFFEY AND ELEANOR MEEHAN
Department of Mathematics, University College
Dublin, Ireland

(Received December 1991)

Abstract—A factorization $x^n I_n = (xI_n - A_1)\cdots(xI_n - A_n)$ where each A_i is an $n \times n$ matrix with minimal polynomial x^n is presented and a uniqueness result is proved for $n = 2, 3$.

1. INTRODUCTION

Let F be an infinite field and let $f(x) \in F[x]$ be a monic polynomial of degree n. In [1], we obtained a factorization

$$f(x)I_n = (xI_n - A_1)\cdots(xI_n - A_n),$$

where $A_1, \cdots, A_n$ are nonderogatory $n \times n$ matrices over F with characteristic polynomial $f(x)$. For ("small") finite fields, the argument presented in [1] fails. In this paper, we obtain a factorization of the above form when $f(x) = x^n$ which is valid over arbitrary fields (and in fact over rings with identity). As a by-product, we obtain a factorization of certain $n \times n$ matrices of determinant one as the product PQ of $n \times n$ matrices P, Q with $P^n = Q^n = I_n$ and we also exhibit a pair of $n \times n$ matrices (X, Y) with X a permutation matrix which has property L and generates the full matrix algebra $M_n(F)$. We also obtain a uniqueness result for the factorization of $x^n I_n$ when $n = 3$.

2. A PRELIMINARY RESULT

Let R be a commutative ring with identity and assume that the matrices occurring in this section have entries in R.

Let $n \geq 3$ be odd,

$$A = \begin{bmatrix} a_1 & y_1 & 0 & \cdots & 0 \\ 0 & a_2 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & y_{n-1} \\ 0 & \cdots & \cdots & 0 & a_n \end{bmatrix}$$

and $P = (p_{ij})$ the permutation matrix corresponding to the n-cycle $(1, 2, 3, \ldots, n)$ (so $p_{i\ i+1} = 1$ $(i = 1, 2, \ldots, n-1)$, $p_{n1} = 1$, $p_{ij} = 0$ otherwise).

Let $B = AP^{-2}$. Let z be an indeterminate.

Using the Laplace expansion along row one, we obtain a recurrence relation for the determinant

$$\det(zI + B) = z^n + a_2 a_n z^2 \Delta(3, \ldots, n-2) - a_1 z \Delta(2, \ldots, n-1) - a_n y_1 z \Delta(2, \ldots, n-2) + a_1 a_2 \ldots a_n$$

where

$$\Delta(k+1,\dots,k+l)=\begin{vmatrix} y_{k+1} & z & 0 & \cdots & \cdots & \cdots & 0 \\ a_{k+2} & y_{k+2} & z & 0 & & & \vdots \\ 0 & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & & \ddots & \ddots & z \\ 0 & \cdots & \cdots & \cdots & 0 & a_{k+l} & y_{k+l} \end{vmatrix}$$

Using the Laplace expansions along the top row and along the bottom row, respectively, we obtain the following two recurrence relations

$$\Delta(k+1,\dots,k+l)=y_{k+1}\Delta(k+2,\dots,k+l)-a_{k+2}z\Delta(k+3,\dots,k+l)$$

and

$$\Delta(k+1,\dots,k+l)=y_{k+l}\Delta(k+1,\dots,k+l-1)-za_{k+l}\Delta(k+1,\dots,k+l-2). \qquad (*)$$

Hence

$$\begin{aligned}\det(zI+B)=z^n+z\big[z\big(a_2a_n\Delta(3,\dots,n-2)+a_1a_{n-1}\Delta(2,\dots,n-3)\big)\\ -(a_1y_{n-1}+a_ny_1)\Delta(2,\dots,n-2)\big]+a_1a_2\cdots a_n.\end{aligned} \qquad (1)$$

Using $(*)$ again we may write (for $n\geq 5$)

$$\begin{aligned}a_2a_n\Delta(3,\dots,n-2)+a_1a_{n-1}\Delta(2,\dots,n-3)=(a_2a_ny_{n-2}+a_1a_{n-1}y_2)\Delta(3,\dots,n-3)\\ -z\big(a_1a_{n-1}a_3\Delta(4,\dots,n-3)+a_2a_na_{n-2}\Delta(3,\dots,n-4)\big).\end{aligned} \qquad (2)$$

We now may use $(*)$ again on the terms

$$a_1a_{n-1}a_3\Delta(4,\dots,n-3)+a_2a_na_{n-2}\Delta(3,\dots,n-4)$$

and continue the process indefinitely.

At each stage, we get an expression of the form $U-Vz$. We then set the term $U=0$ in all these equations. This leads to the following system of equations

$$\begin{aligned} a_ny_1+a_1y_{n-1}&=0\\ a_na_2y_{n-2}+a_1a_{n-1}y_2&=0\\ a_na_2a_{n-2}y_3+a_1a_{n-1}a_3y_{n-3}&=0\\ a_na_2a_{n-2}a_4y_{n-4}+a_1a_{n-1}a_3a_{n-3}y_4&=0\\ a_na_2a_{n-2}a_4a_{n-4}y_5+a_1a_{n-1}a_3a_{n-3}a_5y_{n-5}&=0\\ &\vdots \end{aligned} \qquad (3)$$

The l^{th} equation is as follows for $l=2k+1$:

$$a_na_2a_{n-2}a_4a_{n-4}\cdots a_{n-2k}y_{2k+1}+a_{n-1}a_1a_{n-3}a_3\cdots a_{2k-1}a_{n-(2k-1)}a_{2k+1}y_{n-(2k+1)}=0$$

and for $l=2k$:

$$a_na_2a_{n-2}a_4\cdots a_{2k}y_{n-2k}+a_1a_{n-1}a_3a_{n-3}\cdots a_{2k-1}a_{n-(2k-1)}y_{2k}=0.$$

The last term occurs for $l=(n-1)/2$.

If $y_1,\dots,y_{n-1}$ are chosen to satisfy the system, then we obtain $\det(zI+B)=z^n+a_1a_2\cdots a_n$, for that choice of A.

Suppose $a_1, \ldots, a_n$ are nonzero. Then we can solve the system for $y_1, \ldots, y_{n-1}$. In fact, we may take

$$\begin{aligned} y_1 &= -a_1x_1, & y_{n-1} &= a_nx_1, \\ y_2 &= -a_na_2x_2, & y_{n-2} &= a_1a_{n-1}x_2, \\ y_3 &= -a_1a_3a_{n-1}x_3, & y_{n-3} &= a_na_2a_{n-2}x_3, \quad \text{etc.} \end{aligned}$$

for any elements $x_1, x_2, \ldots, x_{(n-1)/2}$ of R.

Thus there are $(n-1)/2$ "free" parameters and for each choice, the corresponding matrix B has characteristic polynomial $x^n - a_1a_2\cdots a_n$.

A particular case of interest is the case where $a_1, \ldots, a_n$ are equal. Then we can take

$$y_1 = x, \quad y_2 = -x, \quad y_3 = x, \quad y_4 = -x, \quad \ldots, \quad y_{n-1} = -x$$

for any x in R.

Note that if F is a field and $X \in GL(n, F)$ is a nonderogatory matrix with its eigenvalues in F, then X is similar over F to a matrix A of the form above with the y_i nonzero and hence if n is odd, it follows that $X = YZ$ where Z is similar to P and Y has characteristic polynomial $z^n - \det X$. In particular, if $\det X = 1$, then Y is similar to P (and to P^{-1}). [This is clear if the characteristic of F does not divide n since Y has characteristic polynomial $x^n - 1$, while an argument based on the minors of $xI - Y$ (which we omit) yields the result in general.] This gives an explicit version of a special case of a well-known theorem of R. C. Thompson [3] which states that if F is a field and $Q \in M_n(F)$ has determinant one, then $Q = K^{-1}L^{-1}KL$ for some $K, L \in M_n(F)$ except when $(n, |F|) = (2, 2)$. See [2] for a discussion of other approaches to Thompson's theorem.

Choosing indeterminates x, y and substituting

$$a_1 = a_2 = \cdots = a_n = x, \quad y_1 = y, \quad y_2 = -y, \quad \ldots, \quad y_i = (-1)^{i+1}y, \quad \ldots, \quad y_{n-1} = -y,$$

the system of equations (2) is satisfied and it follows that

$$\det(xI + yJ_n + zP^2) = x^n + z^n$$

where $J_n = (w_{ij})$ is the strictly upper-triangular $n \times n$ matrix with its first super-diagonal

$$(w_{12}, w_{23}, \ldots, w_{n-1\,n}) = (1, -1, 1, -1, \ldots, 1, -1),$$

and all other entries zero.

Hence (J_n, P^2) is a pair of matrices with property L [4]. It is easy to check that the pair (J_n, P^2) generate the full matrix algebra $M_n(F)$.

COMMENT. If one attempts to use P itself instead of P^{-2} in the definition of B, then in order to make the process succeed for all fields, the following binomial coefficients must be even

$$\binom{n-1}{1}, \quad \binom{n-2}{3}, \quad \binom{n-3}{4}, \cdots,$$

(where $\binom{k}{l} = 0$ if $k < l$). This occurs if $n = 2^m - 1$ for some integer $m \geq 2$.

3. A FACTORIZATION THEOREM FOR x^nI_n.

Let R be a commutative ring with identity, $n \geq 1$ an odd integer and $M_n(R)$ the ring of $n \times n$ matrices over R. Let x be an indeterminate and let A be defined as in Section 1 with $a_1 = \cdots = a_n = x$, $y_i = (-1)^{i+1}$, $(i = 1, 2, \ldots, n-1)$. So $A = xI_n + J_n$, where J_n is the same as in Section 1. Then

$$\det(xI + AP^{-2}) = z^n + x^n$$

(from Section 1) and hence, using the Cayley-Hamilton theorem,

$$(AP^{-2})^n = x^n I_n,$$

so

$$(xI_n - K_1)\cdots(xI_n - K_n) = x^n I_n,$$

where $K_1 = J_n$ and $K_{i+1} = P^{-2i}K_1P^{2i} > (i = 1, 2, \ldots, n-1)$.

If $n = 2m \geq 2$ is an even integer, then if $J(n)$ is the $n \times n$ Jordan block with characteristic polynomial x^n, there exists a permutation matrix Q with

$$(J(n))^2 = Q^{-1}\begin{pmatrix} J(m) & 0 \\ 0 & J(m) \end{pmatrix} Q.$$

Hence

$$(xI_n - J(n))(xI_n + J(n)) = Q^{-1}\begin{pmatrix} xI_m - J(m) & 0 \\ 0 & xI_m - J(m) \end{pmatrix} Q.$$

A simple inductive argument, using the fact that $J(k)$ and J_k are similar via the diagonal matrix $\operatorname{diag}(1, -1, 1, -1, \ldots, -1, 1)$ when k is odd, yields the following result.

THEOREM 1. *Let R be a commutative ring with identity and let n be a positive integer. Then*

$$x^n I_n = (xI_n - K_1)\cdots(xI_n - K_n),$$

where $K_1, \ldots, K_n \in M_n(R)$ are all similar via elements of $GL(n, R)$ to the full Jordan block corresponding to x^n.

We note that when n is odd, any two of the K_i generate the full matrix algebra $M_n(R)$.

The factorization in Theorem 1 is not unique. Of course, if $n = 2$, the equation $x^2 I_2 = (xI - A_1)(xI - A_2)$ implies $A_2 = -A_1$, so if $x^2 I_2 = (xI - A_1)(xI - A_2)$ with A_1 nonderogatory, then there exists a nonsingular matrix T with $T^{-1}A_1T = K_1$, $T^{-1}A_2T = K_2$. We conclude with a uniqueness result for $n = 3$.

THEOREM 2. *Let F be a field and suppose A_1, A_2 $A_3 \in M_3(F)$ are such that*

$$x^3 I_3 = (xI_3 - A_1)(xI_3 - A_2)(xI_3 - A_3)$$

and trace $(A_1A_2^2) = a^3 \neq 0$. Then there exists $T \in GL(n, F)$ such that

$$T^{-1}A_iT = aK_i, \quad (i = 1, 2, 3).$$

PROOF. Taking determinants, we see that each A_i has characteristic polynomials x^3, so each is nilpotent.

Comparing coefficients yields the equations $A_1 + A_2 + A_3 = 0$, $A_1A_2 + A_1A_3 + A_2A_3 = 0$ and $A_1A_2A_3 = 0$. As a consequence we obtain $A_1A_2 = (A_1 + A_2)^2$, so

$$A_1^2 + A_2A_1 + A_2^2 = 0. \qquad \circledast$$

Thus, since $A_1^3 = 0$, trace $(A_1^2A_2) = -\operatorname{trace}(A_1A_2^2) \neq 0$, so $A_1^2 \neq 0$. Hence, replacing A_1 by $T_1^{-1}A_1T_1$ for some $T_1 \in GL(3, F)$, we may assume $A_1 = aK_1$. The trace condition implies that the $(3, 1)$ entry of A_2 is not zero and hence replacing A_2 by $T_2^{-1}A_2T_2$ for some $T_2 \in GL(3, F)$ of the form

$$T_2 = \begin{bmatrix} 1 & p & q \\ 0 & 1 & -p \\ 0 & 0 & 1 \end{bmatrix}.$$

(Note that T_2 commutes with K_1), we may assume the $(1, 1)$ and $(2, 1)$ entries of A_2 are zero. Thus

$$A_2 = \begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Since trace $(A_2A_1) = 0$, (using $\circledast$), $a_{32} = 0$. Also $A_2^2A_1^2 = 0$, so $a_{13} = a_{23} = 0$. Since A_2 is nilpotent, it now follows that $a_{22} = 0$ and then that $a_{33} = 0$. So

$$A_2 = \begin{pmatrix} 0 & a_{12} & 0 \\ 0 & 0 & 0 \\ a_{31} & 0 & 0 \end{pmatrix}.$$

The equation trace $(A_1A_2^2) = a^2$ now yields $a_{31} = a$ and $(*)$ yields $a_{12} = -a$. Hence $A_2 = aK_2$ and then since $A_1 + A_2 + A_3 = 0$, $A_3 = aK_3$, completing the proof.

References

1. T.J. Laffey and E. Meehan, An extension of a factorization theorem of Wedderburn to matrix rings, *Linear Algebra Appl.* (to appear).
2. T.J. Laffey, Products of matrices, *Generators and Relations in Groups and Geometries,* (Edited by Barlotti, *et al.), NATO ASI Series*, 95–124, Kluwer, (1991).
3. R.C. Thompson, Commutators in the special and general linear groups, *Trans. Amer. Math. Soc.* **101**, 16–33 (1961).
4. T.S. Motzkin and O. Taussky, Pairs of matrices with property L (I), *Trans. Amer. Math. Soc.*, **73**, 108–114, (1952); **80**, 387–401, (1955).